Z. Wahrscheinlichkeitstheorie verw. Gebiete
36, 333 – 343 (1976)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Independence of Events and of Random Variables

Bernard R. Gelbaum
Department of Mathematics, SUNY at Buffalo, N.Y., 4246 Ridge Lea Road, Amherst, N.Y. 14226, USA

The phenomenon of independence of random variables is shown to be singular in that, e.g., there are both finite and infinite sample spaces on which two random variables can be independent iff one is constant. Furthermore on [0, 1], with Lebesgue measure for probability, the usual function spaces contain dense subsets each member of which is independent only of constants. Finally, the requirement of independence among a set of orthonormal functions is shown to imply, in all but trivial instances, that the orthogonal complement of the space is infinite-dimensional.

0. Introduction

The distinguishing feature of probability theory vis-à-vis measure theory is the role played in the former by the concept of *independence*. In the following, independence is studied from the standpoint of its rarity as a phenomenon. The results show that independent sets and independent functions are thinly distributed and, in some instances, (essentially) absent.

1. Independent Events

In correspondence about an ealier draft of this paper, E.O. Thorp remarked that if X is a finite or countably infinite sample space, $X=\{x_1, x_2, \ldots\}$, then the assignment of probabilities: $P(\{x_k\})=2^{-k!}$, $k=2, 3, \ldots$, and $P(\{x_1\})=1-\sum_{k\geqq 2} P(\{x_k\})$, produces a probability measure on the σ-field $\mathscr{S}$ of (all) subsets of X and, with respect to this measure, if $A, B\in\mathscr{S}$ and $\{A, B\}\cap\{\phi, X\}=\phi$ then A and B are not independent.

If X is finite, $X=\{x_1, x_2, \ldots, x_n\}$ then the assignment: $P(\{x_k\})=\varepsilon$, where ε is irrational and $0<\varepsilon<\dfrac{1}{n-1}$, for $k=2, 3, \ldots, n$, $P(\{x_1\})=1-(n-1)\varepsilon$, also yields no nontrivial pairs of independent events.

Finally, if X is countably infinite, $X=\{x_1,x_2,\dots\}$, let $\mathbb{N}$ be decomposed into infinitely many pairwise disjoint infinite subsets $\mathbb{N}_2,\mathbb{N}_3,\dots$ and let $P(\{x_k\})\equiv p_k=\sum_{m\in\mathbb{N}_k}2^{-m!}$, $k=2,3,\dots$, $P(\{x_1\})\equiv p_1=1-\sum_{k\geqq 2}p_k$. Then if $A,B\in\mathcal{S}$, $P(A\setminus B)$, $P(B\setminus A)$ and $P(A\cap B)$ are three numbers p,q,r, each a sum of some p_k. (For purposes of discussing independence, it may be assumed that $A\cup B\not\ni x_1$.) The independence of A and B implies that $(p+r)(q+r)=r$ and in particular that p,q,r are algebraically dependent. However, the arguments proving the transcendence of the classical Liouville numbers show here that p,q,r are always algebraically independent. This last construction is due to S. Schanuel.

Hence there are many kinds of examples of sample spaces that admit no nontrivial pairs of independent events and hence no nontrivial sets of independent random variables. The examples indicated above deal exclusively with atomic sample spaces (from which the early studies in probability sprang). In the next section attention is focussed on the sample space $X\equiv[0,1]$ with Lebesgue measure as probability defined for members of the σ-field $\mathcal{S}$ of Borel-measurable sets. In this setting, the singularity of the condition of independence will appear in a different but equally emphatic way.

2. Random Variables

All random variables considered will be real-valued and defined on $X\equiv[0,1]$ endowed with Lebesgue measure P for the σ-field $\mathcal{S}$ of Borel sets. Since every nonatomic separable sample space is measure-isomorphic to $\{X,\mathcal{S},P\}$, the results below have a generality somewhat greater than their superficial appearance might suggest.

For any random variable f, Ind (f) stands for the set of random variables g such that f and g are independent. If K is the set of constant functions then for all f, $K\subset$ Ind (f). Some of the results below have the reversed conclusion as the central statement. Clearly independence is "stable modulo null sets," i.e., if A and B are independent sets and if M and N are null sets then each of the pairs $\{A\circ M, B\circ N\}$ is independent, where "$\circ$" may mean "$\cup$" or "$\setminus$" for either member or both members of the pair. Thus, as in several other parts of analysis, theorems and proofs below are to be read "modulo null sets". For example, the statement, "Every nonempty measurable set B has metric density 1 everywhere in B," is to be so interpreted.

(2.1) **Lemma.** *Let $f\in C([0,1])$ and for some y_0 assume:* (a) $f^{-1}(y_0)\equiv\{x_1,x_2,\dots,x_n\}$ *is finite of cardinality $n\geqq 1$;* (b) f' *exists and is continuous in some neighborhood of each x_i;* (c) $0\neq|f'(x_i)|=\dfrac{1}{\alpha_i}$.

If B is a nonnull Borel set in $[0,1]$ *and if f and χ_B (the characteristic function of B) are independent, then*

$$P(B)=\frac{\sum_{x_i\in B}\alpha_i}{\sum_{i=1}^{n}\alpha_i}.$$

Proof. Let $\inf_i |f'(x_i)| = 2m$. Then $2m > 0$ and since f' is continuous in neighborhoods of the x_i, there are open intervals U_i around the x_i where f is strictly monotone and $|f'(x)| \geqq m$. On the compact set $[0,1] \setminus \bigcup_{i=1}^{n} U_i$, $|f(x) - y_0| \geqq \delta$ for some $\delta > 0$. Then for ε in $(0, \delta)$, $f^{-1}((y_0 - \varepsilon, y_0 + \varepsilon)) \subset \bigcup_{i=1}^{n} U_i$ and, since f is strictly monotone in U_i, $f^{-1}((y_0 - \varepsilon, y_0 + \varepsilon)) \cap U_i$ is an open interval I_i^ε containing x_i, whence

$$f^{-1}((y_0 - \varepsilon, y_0 + \varepsilon) = \bigcup_{i=1}^{n} I_i^\varepsilon.$$

Clearly, if ε is small enough the intervals I_i^ε are pairwise disjoint. Below I_i^ε is written simply as I_i. Furthermore, owing to the fundamental properties of derivatives, $P(I_i) = (\alpha_i(y_0) + \eta_i(\varepsilon))\, 2\varepsilon$, where $\eta_i(\varepsilon) \to 0$ as $\varepsilon \to 0$. Thus

$$\begin{aligned} P(f^{-1}((y_0 - \varepsilon, y_0 + \varepsilon)) \cap B) &= \sum_{i=1}^{n} P(I_i \cap B) \\ &= P(f^{-1}(y_0 - \varepsilon, y_0 + \varepsilon))\, P(B) \end{aligned}$$

and so

$$\begin{aligned} P(B) &= \frac{\sum_{i=1}^{n} P(I_i \cap B)}{\sum_{i=1}^{n} P(I_i)} = \sum_{i=1}^{n} \frac{P(I_i \cap B)}{P(I_i)} \cdot \frac{P(I_i)}{\sum_{i=1}^{n} P(I_i)} \\ &= \frac{1}{\sum_{i=1}^{n} (\alpha_i(y_0) + \eta_i(\varepsilon))} \cdot \frac{1}{2\varepsilon} \cdot 2\varepsilon \sum_{i=1}^{n} \frac{P(I_i \cap B)}{P(I_i)} (\alpha_i(y_0) + \eta_i(\varepsilon)). \end{aligned}$$

Since we may regard B as having metric density 1 everywhere in B and 0 everywhere off B, we find, as $\varepsilon \to 0$, the right member approaches

$$G(B, f, y_0) \equiv \sum_{x_i \in B} \alpha_i(y_0) \Big/ \sum_{i=1}^{n} \alpha_i(y_0).$$

Note that $G(B, f, y_0)$, for f and y_0 fixed and B variable, is capable of assuming at most 2^n different values. On the other hand, for a nonconstant continuous function g defined on $[0, 1]$, the set $\{P(g^{-1}(A)) \colon A \text{ a Borel set}\}$ is infinite. These remarks lead to

(2.1) **Corollary.** *For f as described in Lemma* 2.1, $\mathrm{Ind}(f) \cap C([0,1]) = K$.

(2.1) *Example.* Let $f(x) = \sin 2\pi x$, $0 < y_0 < 1$, or $-1 < y_0 < 0$, $0 < P(B) < 1$. Then the numerator and denominator of $G(B, f, y_0)$ are respectively $2\pi\sqrt{1 - y_0^2}$ and $4\pi\sqrt{1 - y_0^2}$ whence $G(B, f, y_0) = \frac{1}{2}$. A similar calculation and Corollary 2.1 lead to the observation:

$$\left[\bigcup_{k=1}^{\infty} (\mathrm{Ind}(\cos 2\pi k x) \cup \mathrm{Ind}(\sin 2\pi k x))\right] \cap C([0,1]) = K.$$

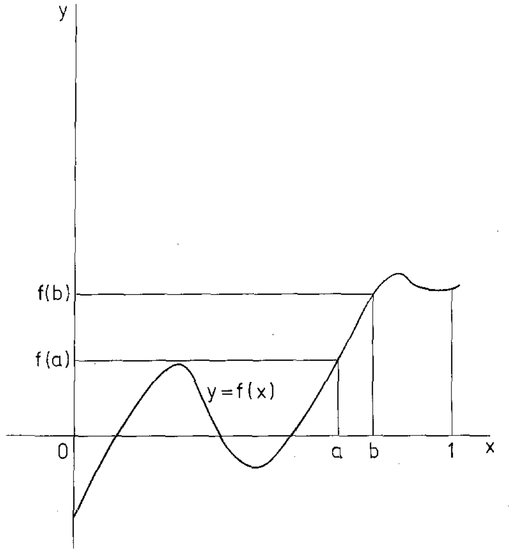

Fig. 2.1

The example suggests that for a fairly general and simply described set of functions f, $\mathrm{Ind}(f) \cap C([0,1]) = K$. Indeed, there obtains

(2.1) **Theorem.** *If f is the restriction to $[0,1]$ of a function $\tilde{f}$ analytic in a region containing $[0,1]$ and if f is not constant then* $\mathrm{Ind}(f) \cap C[0,1] = K$.

Proof. The identity theorem for analytic functions shows that for each y, $\mathrm{card}(f^{-1}(y))$ is finite.

If for each y there is an $x_y \in f^{-1}(y)$ where $f'(x_y)=0$, then since $f^{-1}(y_1) \cap f^{-1}(y_2) = \phi$ if $y_1 \neq y_2$, the set $\{x_y\}$ is infinite and has a limit point x_0 in $[0,1]$. By the identity theorem for analytic functions, $f' \equiv 0$, f is constant, a contradiction. Thus for some y_0, $f' \neq 0$ on $f^{-1}(y_0)$ and the function $G(B, f, y_0)$ is definable. By Corollary 2.1 the result follows.

(2.2) *Example.* Let f be measurable. Assume that for some subinterval $[a,b]$, $f^{-1}(f([a,b])) = [a,b]$ and that f^{-1} is measurable on $f([a,b])$. Then $\mathrm{Ind}(f) = \mathrm{K}$. The basic property of f is shown in Figure 2.1. (Continuity of f is *not* assumed.) To show $\mathrm{Ind}(f) = K$ the following development serves.

For any measurable set $S \subset [0,1]$, let

$$\varphi_S(x) = \begin{cases} x, & x \in S \\ 0, & x \notin S \end{cases}.$$

Then there obtains

(2.2) **Lemma.** *For* $0 \leqq a < b \leqq 1$, $\mathrm{Ind}(\varphi_{[a,b]}) = K$.

Proof. If a measurable $g \notin K$, then for some Borel set A, $0 < P(g^{-1}(A)) < 1$. On the other hand if $g \in \mathrm{Ind}(\varphi_{[a,b]})$, then for all Borel sets, in particular for A, and $a \leqq c < d \leqq b$

$$P(g^{-1}(A)) = \frac{P(g^{-1}(A) \cap [c,d])}{P([c,d])}$$

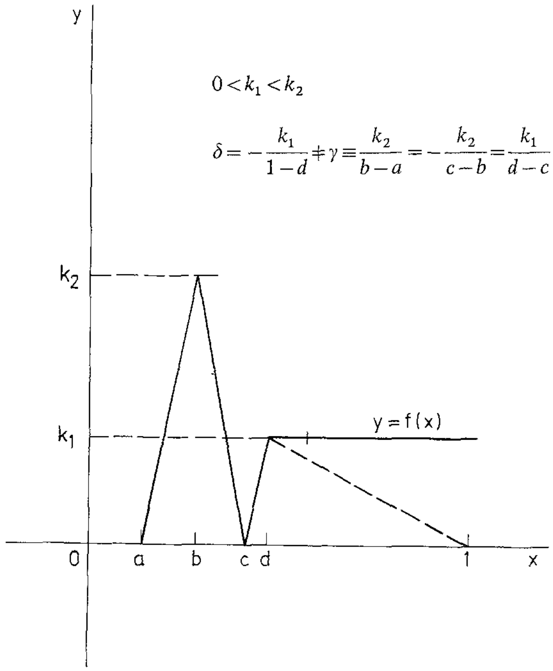

Fig. 2.2

since $[c, d]=\varphi_{[a,b]}^{-1}([c, d])$. By metric density arguments, as $|d-c|\to 0$, the right member approaches either 0 or 1 whereas the left member is in $(0, 1)$, a contradiction.

The assertion of Example 2.2 is a corollary to Lemma 2.2.

Indeed, if $g\in \mathrm{Ind}(f)$ let

$$k(y)=\begin{cases} f^{-1}(y), & y\in f([a, b]) \\ 0, & \text{otherwise} \end{cases}$$

and let $h=k\circ f$. Then $h=\varphi_{[a,b]}$. As a Borel measurable function of the Borel measurable function f, h is also independent of g, a contradiction of Lemma 2.2.

(2.3) *Example.* Let f be the function depicted by the solid line graph in Figure 2.2.

Let $J=f^{-1}([\alpha, \beta])$ and let the Borel set B and J be independent. Then for $0<\alpha<\beta<k_1$, J is the union of three equally long and disjoint subintervals I_1 of $[a, b]$, I_2 of $[b, c]$ and I_3 of $[c, 1]$ and

$$P(J)=\frac{3(\beta-\alpha)}{\gamma}=3\,P(I_j), \quad j=1, 2, 3.$$

Thus

$$\begin{aligned} P(B\cap J)&=P(B\cap I_1)+P(B\cap I_2)+P(B\cap I_3) \\ &=P(B)\,P(J), \end{aligned}$$

and so

$$P(B)=\frac{P(B\cap I_1)}{3\,P(I_1)}+\frac{P(B\cap I_2)}{3\,P(I_2)}+\frac{P(B\cap I_3)}{3\,P(I_3)}.$$

If $|\beta-\alpha|\to 0$, $0<\alpha<\beta<k_1$, metric density arguments now show that $P(B)=0, \frac{1}{3}, \frac{2}{3}$, or 1. On the other hand, if $k_1<\alpha<\beta<k_2$, then similar calculations show that $P(B)=0, \frac{1}{2}$ or 1. Hence the only Borel sets B independent of all sets $f^{-1}(J)$ are $[0, 1]$ and ϕ and so $\operatorname{Ind}(f)=K$.

If the graph of f is as in Figure 2.2 for $0\leqq x\leqq c$ and then follows the dashed line for $c\leqq x\leqq 1$, metric density arguments show that for B to be independent of $f^{-1}(J)$ there must obtain $P(B)=0, \frac{1}{2}, 1$ if $k_1<\alpha<\beta<k_2$; similarly if $0<\alpha<\beta<k_1$, then $P(B)=0, \dfrac{\delta+\varepsilon\delta}{3\delta+\gamma}, \dfrac{2\delta+\varepsilon\gamma}{3\gamma+\delta}, \dfrac{3\delta+\varepsilon\gamma}{3\delta+\delta}$ or $\dfrac{\gamma}{3\gamma+\delta}$, where $\varepsilon=0$ or 1. For $\gamma=1$, $\delta=2$, these values are $0, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}$ or 1. The nontrivial values are all different from the nontrivial value $\frac{1}{2}$. Thus again $\operatorname{Ind}(f)=K$.

Since k_1 and k_2 are arbitrary, it is clear that for any polynomial g, an arbitrarily small portion of the graph of g can be excised and replaced by a graph resembling one of the graphs in Figure 2.2. The resulting function $\tilde{g}$ is continuous and $\operatorname{Ind}(\tilde{g})=K$. Thus there obtains

(2.2) **Theorem.** *For f in $E\equiv L_p$, $1\leqq p<\infty$ or $C([0, 1])$, and $\varepsilon>0$ there is in E an $\tilde{f}$ such that $\|f-\tilde{f}\|_E<\varepsilon$ and* $\operatorname{Ind}(\tilde{f})=K$.

The functions of the types given above can, by "rounding the corners of their graphs," be made to belong to $C^\infty([0, 1])$ without disturbing the validity of the proofs given for the properties of their sets $\operatorname{Ind}(f)$. It is sufficient for such purposes, that the corner-rounding be carried out over a union of intervals of total measure less than one.

A routine calculation shows that $\operatorname{Ind}(\sin 2\pi x)$ is $\{c_1\chi_{[\frac{1}{4},\frac{3}{4}]}+c_2\chi_{[0,\frac{1}{4}]\cup[\frac{3}{4},1]}: c_1, c_2\in\mathbb{R}\}$. Further analysis shows that for a wide class of functions f in $C^1([0, 1])$, $\operatorname{Ind}(f)=\{\sum c_k\chi_{S_k}: c_k\in\mathbb{R},\ S_k$ a union of nonoverlapping intervals, $S_k\cap S_l=\phi,\ k\neq l\}$. Precise hypotheses on f for which the above obtains are as follows:

$$f\in C^1([0, 1]); \quad m=\min f<\max f\equiv M; \quad f^{-1}(m)=\{p_i\};$$

$f^{-1}(M)=\{q_j\}$ where $p_1<q_1<p_2<q_2<\dots$; $f'^{-1}(0)=\{p_1, p_2, \dots, q_1, q_2, \dots\}$; for $m<y<M$, $f^{-1}(y)$ is finite, say $\{r_1(y), \dots, r_{n(y)}\}$; $f'(r_k(y))=\dfrac{1}{a_k g(y)}$ where $g(y)>0$ and $g(y)$ is integrable on $[m, M]$.

The situation is exemplified by the functions $\sin 2\pi n x$, $\cos 2\pi n x$ and more generally by functions corresponding to the graph in Figure 2.3. Let $I_1=[p_1, q_1]$, $I_2=[q_1, p_2]$, $I_3=[p_2, q_2]$, ... If $G=\int_m^M g(y)\,dy$, then $a_1=\dfrac{P(I_1)}{G}$, $a_2=\dfrac{P(I_2)}{G}$, etc. Metric density arguments show that if B is independent for $f^{-1}((-\infty, y))$ for all y then $P(B)=\dfrac{\sum_{k\in S} a_k}{\sum_k a_k}$, where $S=\{k: P(B\cap I_k)\neq 0\}$. Furthermore, for k in S, $P(B\cap I_k)=\dfrac{a_k}{\sum_j a_j}=P(I_k)$ whence each such B is the union of such I_k and the conclusion follows.

The above results as well as some below reveal an unusual situation: *a function f for which* $\operatorname{Ind}(f)$ *is a linear set.* The classic example of S. Bernstein shows that this phenomenon is not universal: Let $X=\{x_1, x_2, x_3, x_4\}$, $P(x_i)=\frac{1}{4}$, $i=1, 2, 3, 4$,

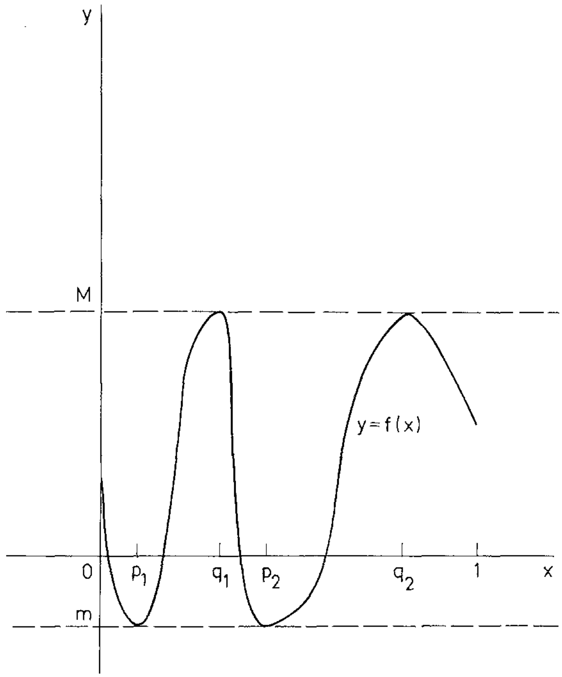

Fig. 2.3

$f_1=\chi_{\{x_1,x_2\}}$, $f_2=\chi_{\{x_1,x_3\}}$, $f_3=\chi_{\{x_1,x_4\}}$. Then $f_2, f_3 \in \operatorname{Ind}(f_1)$ but $f_2-f_3 \notin \operatorname{Ind}(f_1)$ since $(f_2-f_3)^{-1}(1)=\{x_3\}$, $f_1^{-1}(1)=\{x_1,x_2\}$ and $(f_2-f_3)^{-1}(1) \cap f_1^{-1}(1)=\phi$.

If the members of $\operatorname{Ind}(f)$ are *mutually* independent, then $\operatorname{Ind}(f)$ is linear. However, in all the examples given $\operatorname{Ind}(f)$ contains members that are not mutually independent. Nevertheless for these examples $\operatorname{Ind}(f)$ is linear.

(2.3) *Example.* In the following there is given a rather general construction of a continuous, piecewise linear function f such that $\operatorname{Ind}(f)$ is the *linear* set consisting of the linear combinations $\sum_{m=n}^{n-1} c_m \chi_{I_m}$ of the characteristic functions χ_{I_m} of the intervals I_m, $m=0,1,\ldots,n-1$, on which f is linear. The construction obtains for arbitrary $n \in \mathbb{N}$.

(2.3) **Lemma.** *Let $\alpha_0, \alpha_2, \ldots, \alpha_{n-1}>0$ be arbitrary. There are numbers $k>0$ and $0=a_0<a_1<\cdots<a_n=1$ and a continuous piecewise linear function f such that $f(a_0)=f(a_2)=\cdots=0$, $f(a_1)=f(a_2)=\cdots=k$, $f'(x)=(-1)^m\alpha_m$ in (a_m, a_{m+1}). (Two such functions are shown in Fig. 2.4.)*

Proof. Let $k=\left(\sum_{j=1}^{n-1}\frac{1}{\alpha_j}\right)^{-1}$, $a_m=k\left(\sum_{j=1}^{m-1}\frac{1}{\alpha_j}\right)$, $m=1,2,\ldots,n-1$. Then direct calculation verifies that the corresponding f satisfies the requirements given in the statement of the lemma.

(2.4) **Lemma.** *For arbitrary n, there are n positive numbers $\beta_0, \beta_1, \ldots, \beta_{n-1}$ such that if S' and S'' are different subsets of $\{0,1,\ldots,n-1\}$ then $\sum_{S'}\beta_j \neq \sum_{S''}\beta_j$ and $\sum_{m=0}^{n-1}\beta_m=1$.*

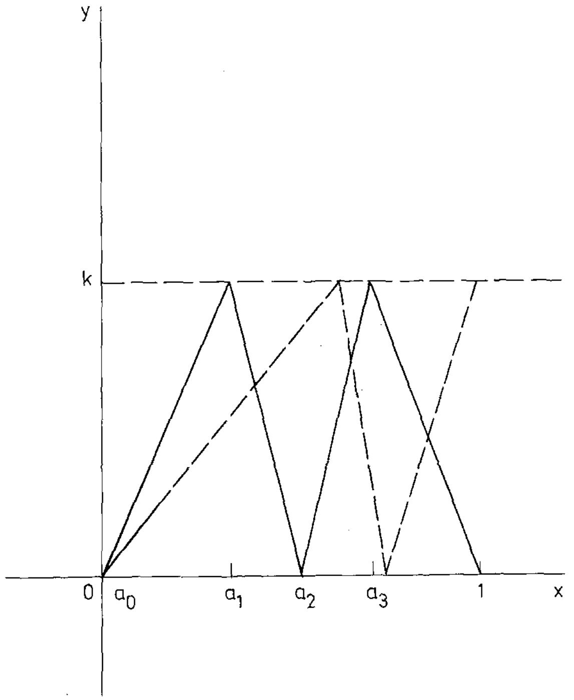

Fig. 2.4

Proof. For each of the $(2^n-1)^2$ pairs S', S'' of subsets of $\{0, 1, \ldots, n-1\}$ let $\Pi(S', S'')$ be the hyperplane defined by the equation $\sum_{S'} x_j - \sum_{S''} x_j = 0$. Let σ be the *open simplex* defined by the conditions $\sum_{j=0}^{n-1} x_j = 1$, $x_j > 0$, $j = 1, 2, \ldots, n-1$. Then $\sigma \setminus \bigcup_{(S', S'')} \Pi(S', S'')$ is nonempty and for $(\beta_0, \beta_1, \ldots, \beta_{n-1})$ in it, the requirements of lemma are satisfied.

(2.3) **Theorem.** *Let $\beta_0, \beta_1, \ldots, \beta_{n-1}$ be chosen according to Lemma 2.4, let $\alpha_m = 1/\beta_m$, $m = 0, 1, \ldots, n-1$, and let $k, a_1, \ldots, a_{n-1}$ be chosen according to Lemma 2.3. Then the corresponding f is such that* $\mathrm{Ind}(f) = \{\sum c_m \chi_{I_m}\}$ *where χ_{I_m} is the characteristic function of the interval $I_m = [a_m, a_{m+1}]$, $m = 0, 1, \ldots, n-1$.*

Proof. It will be shown that a Borel set B is independent of $f^{-1}((\alpha, \beta))$ for all α, β iff B is the union $\bigcup_{m \in S} I_m$, S a subset of $\{0, 1, 2, \ldots, n-1\}$. First, for any α, β such that $f^{-1}((\alpha, \beta)) \neq \phi$, $f^{-1}((\alpha, \beta)) = \bigcup_{n-1}^{n-1} J_m$, where the interval $J_m \subset I_m$ and $P(J_m) = \dfrac{\beta - \alpha}{\alpha_m}$, $m = 0, 1, \ldots, n-1$, whence, $P(f^{-1}((\alpha, \beta))) = \sum_{m=0}^{n-1} \dfrac{\beta - \alpha}{\alpha_m} = (\beta - \alpha) \sum_{m=0}^{n-1} \beta_m = \beta - \alpha$. If $B = \bigcup_{m \in S'} I_m$, then $P(B) = \sum_{m \in S'} P(I_m) = k(\sum_{m \in S'} \beta_m)$ and $P(B \cap f^{-1}((\alpha, \beta))) = \sum_{m \in S'} (\beta - \alpha) \beta_m = k \sum_{m \in S'} \beta_m \cdot \dfrac{\beta - \alpha}{k} = P(B) \cdot \dfrac{\beta - \alpha}{k} = P(B) \cdot P(f^{-1}((\alpha, \beta)))$ since $\dfrac{1}{k} = \sum_{m=0}^{n-1} \beta_m = 1$. Hence B and $f^{-1}((\alpha, \beta))$ are independent.

Conversely, if B and $f^{-1}((\alpha,\beta))$ are independent for all α,β, and if B is not a union $\bigcup_{m\in S} I_m$, $S\subset\{0,1,\dots,n-1\}$, then for some j_0 $P(I_{j_0}\cap B)P(I_{j_0}\setminus B)>0$. As in earlier metric density calculations

$$P(B)=\frac{\sum_{j=0}^{n-1}P(J_j\cap B)}{(\beta-\alpha)}=\sum_{j=0}^{n-1}\frac{P(J_j\cap B)}{P(J_j)}\cdot\frac{(\beta-\alpha)\beta_j}{(\beta-\alpha)}=\sum_{j=0}^{n-1}\frac{P(J_j\cap B)}{P(J_j)}\cdot\beta_j.$$

If $x'\in I_{j_0}$ is a point of metric density 1 for B, (such an x' exists since $P(I_{j_0}\cap B)>0$) and if $\alpha<f(x')<\beta$ and $|\beta-\alpha|\to 0$, then the above shows

$$P(B)=\Big(\sum_{j\in S'}\beta_j\Big)$$

where $j_0\in S'$. On the other hand, if $x''\in I_{j_0}$ is a point of metric density 0 for B (such an x'' exists since $P(I_{j_0}\setminus B)>0$) then the same argument and calculation show

$$P(B)=\Big(\sum_{j\in S''}\beta_j\Big)$$

where $j_0\notin S''$. Thus $S'\neq S''$ and, owing to the way the β_j were chosen, $\sum_{j\notin S'}\beta_j\neq\sum_{j\in S''}\beta_j$. The contradiction shows $B=\bigcup_{m\in S} I_m$, $S\subset\{0,1,\dots,n-1\}$.

If $g\in\operatorname{Ind}(f)$ then for all y, $g^{-1}((-\infty,y))$ is a Borel set B and so $g^{-1}((-\infty,y))=\bigcup_{m\in S} I_m$. Clearly $g=\sum_{m=0}^{n-1}c_m\chi_{I_m}$. By the above $g\in\operatorname{Ind}(f)$ iff $g=\sum_{m=0}^{n-1}c_m\chi_{I_m}$.

Remarks. 1. For $0=a_0<a_1<\cdots<a_n=1$ given, the numbers $\alpha_m\equiv(k(a_{m+1}-a_m))^{-1}$, $m=0,1,2,\dots,n-1$, for arbitrary positive k, permit the construction of Lemma 2.3. However, the condition of Lemma 2.4 is not satisfied unless the a_m satisfy $\sum_{m\in S'}(a_{m+1}-a_m)\neq\sum_{m\in S''}(a_{m+1}-a_m)$ for every pair of distinct subsets S', S'' of $\{0,1,\dots,n-1\}$.

2. In the notations of these examples, if the condition $f(a_1)=f(a_3)=\cdots=k$ is relaxed and is replaced by $f(a_1)=k_1$, $f(a_3)=k_3,\dots,$ where $k_1,k_3,\dots>0$, and not all k_j are equal, and if, via, e.g., the Schanuel construction of § 1, the numbers β_m are chosen so that the set of all sums $\{\sum_{m\in S}\beta_m\}$, $S\subset\{0,1,2,\dots,n-1\}$ is algebraically independent over $\mathbb{Z}$, then the corresponding f is such that $\operatorname{Ind}(f)=K$. The existence of such an f follows from the solution of a simple set of equations for $k_1,k_3,\dots,a_1,a_2,\dots,a_{n-1}$ with the boundary condition that not all k_j are equal.

The crucial step is the proof, via metric density arguments, that for any Borel set B independent of $f^{-1}((\alpha,\beta))$ for all α,β, $P(B)$ is given by at least *two* expressions $\dfrac{\sum_{j\in S_1'}\beta_j}{\sum_{j\in S_1''\supset S_1'}\beta_j}$ and $\dfrac{\sum_{j\in S_2'}\beta_j}{\sum_{j\in S_2''\supset S_2'}\beta_j}$ where $S_1''\neq S_2''$, in denial of the assumption of algebraic independence for the sums $\sum_{m\in S}\beta_m$.

3. The denseness in several function spaces of the set of functions f for which $\operatorname{Ind}(f)=K$ shows in a rather extreme way how unnecessarily strong the "mutual independence" hypothesis is in the classical limit theorems of probability. Indeed, for example, if $\{f_n\}$ is a set of mutually independent random variables for which the

central limit theorem holds, then for a suitable norm $\|\ldots\|$ and a suitable sequence of positive numbers $\varepsilon_n \to 0$, there is a sequence $\{\tilde{f}_n\}$ such that $\|f_n - \tilde{f}_n\| < \varepsilon_n$ and for which the conclusion of the central limit theorem holds. However, for each n, $\mathrm{Ind}(\tilde{f}_n) = K$. In other words the $\tilde{f}_n$ are not even *pairwise* independent, much less *mutually* independent.

Let $d(g)$ stand for the maximal number of linearly independent members of $\mathrm{Ind}(g)$, then since $K \subseteqq \mathrm{Ind}(g)$, $d(g) \geqq 1$, and the results above show that for arbitrary $n \in \mathbb{N}$, there is an f_n such that $d(f_n) = n$. Thus in some Banach function space E e.g., $C([0,1])$, let $E_n \equiv \{f : d(f) = n\} = d^{-1}(n) \cap E$, $n = 1, 2, \ldots, \infty$. Then these mutually disjoint sets exhaust E, whence at least one is of the second category. (See § 4.)

3. Independence and Orthogonality

In this section, the underlying sample space X is *arbitrary*. If $\{f_\lambda\}$ is a set of nonconstant mutually independent functions in L^2 then for $a_\lambda \equiv \int f_\lambda$, $\{f_\lambda - a_\lambda\}$ is a set of mutually independent, nonzero and orthogonal functions. Clearly the function 1 is in $\{f_\lambda - a_\lambda\}^\perp$ and so the set $S = \left\{\frac{f_\lambda - a_\lambda}{\|f_\lambda - a_\lambda\|_2}\right\} \cup \{1\}$ is orthonormal and independent. There arises the question: Can $\{f_\lambda\}$ or S or $\{f_\lambda\} \cup \{1\}$ span L^2? The answer is essentially, "no," as the following lines show.

(3.1) **Theorem.** *If $\{f_\lambda\}$ is a set of orthonormal and mutually independent functions in L^2, then for* $\mathrm{card}\{f_\lambda\} \geqq 3$ *the orthogonal complement $M^\perp$ of the span M of the f_λ is itself different from $\{0\}$ and is infinite-dimensional if L^2 is infinite-dimensional.*

Proof. The essential tool is the well-known equation

$$\int_X f_{\lambda_1}^{p_1} f_{\lambda_2}^{p_2} \ldots f_{\lambda_k}^{p_k}\, dP = \prod_{i=1}^{k} \Big(\int_X f_{\lambda_i}^{p_i}\, dP\Big)$$

valid if the f_{λ_i} are mutually different and if the p_i are nonnegative integers. From this equation and the assumed orthogonality relations it follows that for $\lambda \neq \mu$

$$0 = \int_X f_\lambda f_\mu\, dP = \int_X f_\lambda\, dP \int_X f_\mu\, dP.$$

Hence, if for some $\bar{\lambda}$, $\int_X f_{\bar{\lambda}} P \neq 0$, then for all $\lambda \neq \bar{\lambda}$, $\int_X f_\lambda\, dP = 0$, i.e., $\int_X f_\lambda\, dP = 0$ for all λ save at most one $(\bar{\lambda})$. Clearly it may be assumed that each f_λ is a nonnull member of L^2, i.e., $\int_X f_\lambda^2\, dP > 0$. Thus, if $f_{\lambda_1}, f_{\lambda_2}, \ldots, f_{\lambda_k}$ are pairwise different then

$$\int_X f_{\lambda_1}^2 f_{\lambda_2}^2 \ldots f_{\lambda_k}^2\, dP = \prod_{i=1}^{k} \Big(\int_X f_{\lambda_i}^2\, dP\Big) > 0$$

and so all products $f_{\lambda_1} f_{\lambda_2} \ldots f_{\lambda_k}$ of pairwise different members of $\{f_\lambda\}$ are nonnull members of L^2.

Let $f_{\bar{\lambda}}, f_{\lambda_1}, f_{\lambda_2}, \ldots, f_{\lambda_k}$ be pairwise different. Then for $k \geqq 2$, $f_{\lambda_1} f_{\lambda_2} \ldots f_{\lambda_k}$ is a (nonnull) member of $M^\perp$. Indeed, for $f_{\lambda_0} \in \{f_\lambda\}$,

$$\int_X (f_{\lambda_1} f_{\lambda_2} \ldots f_{\lambda_k}) f_{\lambda_0}\, dP = \prod_{i=0}^{k} \Big(\int_X f_{\lambda_i}\, dP\Big) \quad \text{if } f_{\lambda_0}, f_{\lambda_1}, \ldots, f_{\lambda_k}$$

are pairwise different,

$$= \prod_{i \neq j} \left(\int_X f_{\lambda_i} dP \right) \int_X f_{\lambda_j}^2 dP \qquad \text{if } f_{\lambda_0} = f_{\lambda_j}.$$

In either event, since $k \geqq 2$ and $f_{\lambda_i} \neq f_{\bar{\lambda}}$ it follows that $f_{\lambda_1} f_{\lambda_2} \ldots f_{\lambda_k}$ is in $M^\perp$.

If $f_{\bar{\lambda}}, f_{\lambda_1}, \ldots, f_{\lambda_k}$ are pairwise different and if $f_{\bar{\lambda}}, f_{\mu_1}, f_{\mu_2}, \ldots, f_{\mu_l}$ are pairwise different and if $f_{\lambda_1} f_{\lambda_2} \ldots f_{\lambda_k} = f_{\mu_1} f_{\mu_2} \ldots f_{\mu_l}$ then $k = l$ and $\{f_{\lambda_1}, f_{\lambda_2}, \ldots, f_{\lambda_k}\} = \{f_{\mu_1}, f_{\mu_2}, \ldots, f_{\mu_l}\}$. Indeed, multiplying both sides of the last numerical equation preceding by $f_{\lambda_1} f_{\lambda_2} \ldots f_{\lambda_k}$ and integrating leads to

$$0 < \int_X f_{\lambda_1}^2 f_{\lambda_2}^2 \ldots f_{\lambda_k}^2 dP = \int_X f_{\lambda_1} f_{\lambda_2} \ldots f_{\lambda_k} f_{\mu_1} f_{\mu_2} \ldots f_{\mu_l} dP.$$

If $\{f_{\lambda_1}, f_{\lambda_2}, \ldots, f_{\lambda_k}\} \neq \{f_{\mu_1}, f_{\mu_2}, \ldots, f_{\mu_l}\}$, the right side of the above may be written $\left(\int_X F\,dP\right)\left(\int_X f_? dP\right)$ where $f_?$ is in the symmetric difference of $\{f_{\lambda_1}, f_{\lambda_2}, \ldots, f_{\lambda_k}\}$ and $\{f_{\mu_1}, f_{\mu_2}, \ldots, f_{\mu_l}\}$. However, $\int_X f_? dP = 0$ and a contradiction obtains.

Thus the set of W of products

$$\{f_{\lambda_1} f_{\lambda_2} \ldots f_{\lambda_k} : f_{\bar{\lambda}}, f_{\lambda_1}, f_{\lambda_2}, \ldots, f_{\lambda_k} \text{ pairwise different, } k \geqq 2\}$$

is contained in $M^\perp$.

The arguments show that different sets $\{f_{\bar{\lambda}}, f_{\lambda_1}, f_{\lambda_2}, \ldots, f_{\lambda_k}\}$ generate different products $f_{\lambda_1} f_{\lambda_2} \ldots f_{\lambda_k}$ and that these are pairwise orthogonal. Thus if $\{f_\lambda\}$ is infinite, $M^\perp$ is infinite-dimensional. When $\{f_\lambda\}$ is finite, say cardinality $(\{f_\lambda\}) = n+1$ then clearly the dimension of $M^\perp$ is at least $\binom{n}{2} + \binom{n}{3} + \cdots + \binom{n}{n} = 2^n - n - 1$.

Remarks. 1. If X is a sample space consisting of two elements x_1, x_2, let $f_1 \equiv 1$, $f_2(x_1) = -f_2(x_2) = 1$. Then $\{f_1, f_2\}$ consists of orthogonal independent functions if $P(x_1) = P(x_2) = \frac{1}{2}$. Indeed, $\int_X f_1 f_2 dP = 1 \cdot 1 \cdot \frac{1}{2} + 1 \cdot (-1) \cdot \frac{1}{2} = 0$. Furthermore $f_1 \in K$ whence $f_2 \in \operatorname{Ind}(f_1)$. Clearly L^2 is two-dimensional, whence in this case $M^\perp = \{0\}$.

Thus the number *three* in the statement of the theorem is best possible.

2. The construction in the proof is similar to that used to construct the Walsh functions from the Rademacher functions.

If $\{f_\lambda\}$ spans L^2 so does S, which cannot be true by Theorem 3.1.

4. Problem

1. For the usual Banach function spaces E on $[0, 1]$ which of the sets $d^{-1}(n) \cap E$ is (are) of the second category?

Reference

1. Renyi, A.: Foundations of probability, p. 129. San Francisco: Holden-Day 1970

Received December 17, 1975; in revised form May 30, 1976